Rapid Data Collection using Automated Model Generation and Performance Evaluation

A workflow for morphological studies of apartment floor plans

Sven Schneider¹, Martin Bielik², Dirk Donath³, Michel Triemer⁴, Julia Tschetwertak⁵, Alexander Hollberg⁶
¹,²,³,⁴,⁵ Bauhaus-Universität Weimar. ⁶ Bauhaus-University Weimar
¹,²,³,⁴,⁵,⁶ {sven.schneider|martin.bielik|dirk.donath|michel.triemer|julia.tschetwertak|alexander.hollberg}@uni-weimar.de

In this paper we propose a framework for accelerating the process of data-collection from apartment floor plans. After describing the general workflow and the criteria used for performance evaluation, we present a tool for automated model generation and evaluation. To demonstrate the functionality of this workflow we show examples, based on a preliminary test phase undertaken in a seminar for master students.

Keywords: Data Collection, Apartment Floor Plans, Space Syntax, Automated Model Generation, Automated Performance Evaluation

INTRODUCTION

Morphological studies are a central topic of architectural research. The objective of these studies is to find correlations between design variables and performance criteria. Design variables refer to the metric and topological properties of geometric elements. This, for example, can be the number, position and dimension of rooms, walls, doors or windows and their relationships to each other. Performance criteria refer to measures derived quantitatively from a certain configuration of these design variables. One can differentiate between criteria that can be directly calculated from a configuration (such as the amount of daylight in or the visual centrality of a room), and those that can only be empirically derived (all kinds of human behaviour). Knowledge of the relationships between design variables and performance criteria is central to the process of designing, since it enables one to define design variables more purposefully. So far several studies already exist that focus on correlations between design variables (Steadman, 1983), correlations between directly measurable performance criteria (Abshirini & Koch, 2013; Güney, 2007; Hanson, 1998; Hillier, Hanson, & Graham, 1987; Peponis, 2012) or correlations between directly measurable performance criteria and human behaviour (Franz & Wiener, 2008; Hillier & lida, 2005; Sevtuk, 2010). To obtain meaningful results, it is necessary to sample a large number of cases and to consider as many criteria as possible. The latter is motivated by the fact that one cannot know in advance which criteria are most important for certain research questions as well as an awareness that the complexity of the environment and the human cognitive process is such that the majority of phenomena we are examining is the result of multiple variables acting simul-
taneously. Using advanced statistical methods with a large number of variables places high demands on the sample size (Tanaka, 1987).

The basic requirement for any data analysis is firstly data collection and secondly raw data processing to clear and prepare the data depending on the particular analysis method. However, the collection of a large number of datasets with many performance criteria is very time consuming. In particular, the preparation of models and analysis of these models requires various programs (e.g. drawing/modeling in AutoCAD, spatial analysis in DepthMap, daylight analysis in Daysim). In this paper we propose a framework for accelerating the process of data-collection from apartment floor plans. After describing the general workflow and the criteria used for performance evaluation, we present a tool for automated model generation and evaluation. To demonstrate the functionality of this workflow we show examples, based on a preliminary test phase undertaken in a seminar for master students.

MODELS FOR APARTMENT FLOOR PLAN ANALYSIS

The framework presented in this paper arose in the context of a study project concerned with the morphology of apartment buildings. Apartment buildings represent one of the most common building types worldwide. Consequently, this type of building has a high economic and ecological impact, and is therefore a relevant example for ascertaining a better understanding of the relationships between design variables and a multitude of performance criteria (social, economic and ecological). With regarding to design variables, the main focus of our project is on the geometry of space. This focus is due to the fact that the definition of the geometry is one of the most important steps in the design process because, compared to other variables (e.g. the material or color of surfaces), the geometry can't be changed that easily after the building has been erected.

A crucial step in order to be able to adequately investigate the influence of floor plan geometry on a range of performance criteria is the modelling of apartments. Generally speaking, apartment needs to be analysed at multiple scales. This is because the global characteristics of an apartment give us little indication of how an apartment is experienced in a single room, or from a single point. Furthermore the shape of an apartment is often dictated by its position in a building, and its relationship to the urban surroundings. For each of these scales, we can determine measures for several performance criteria. In this paper we will concentrate on the two smallest scales: the apartment and its individual rooms (see figure 1).

Figure 1
Different levels of scale (apartment, physically closed spaces, convex spaces, analysis points)

Special care needs to be taken concerning the representation of space, because several measures derived later on in the process are based on the representation. Since most of these measures will later be correlated with different kinds of human behaviour it is necessary to find representations suitable for describing spatial properties that relate to human spatial experience. The entity "physically closed spaces", which refers to spaces bounded by walls and doors, is too limited when describing such properties. This becomes obvious when one considers non-convex rooms such as L-shaped rooms or completely open plans: although these rooms can be described as physically closed spaces they are experienced differently from different positions. We therefore decided to use two more representations for discretising the space of an apartment into single elements rather than just modelling the shape of physically closed rooms. The first representation is a convex space (Hillier & Hanson, 1984) while the second is a regu-
lar grid of points filling the entire space of the floor plan (in the following these points are referred to as Analysis Points). Convex Spaces are spaces in a floor plan in which all points within the space are mutually visible. Theoretically there are an infinite number of convex spaces in an environment. To identify an analysable set of convex spaces (convex map) the largest and most compact spaces are always used. In spaces with complex shapes, the creation of this convex map is sometimes ambiguous. However, for most interior spatial configurations it serves as a useful tool that has been used for many morphological studies.

Physically closed rooms as well as convex spaces are still a very approximate way of discretizing space into single elements. The use of analysis points makes it possible to take measurements for each position in space. Furthermore it has the advantage that the generation of points is unambiguous. To link these different scales we created a database consisting of tables for each scale, whereby every apartment contains a list of IDs to single rooms / convex spaces and analysis points and vice versa. This mutual links make it possible to use measurements of the analysis points to derive global characteristics for larger scales, for example to calculate global daylight characteristics (such as average, minimum, maximum, standard deviation) for each convex space, physically closed room and entire apartment the from the daylight measurements of the analysis points.

**PERFORMANCE CRITERIA**

As stated in the introduction, morphological studies can have different purposes: to examine correlations between design variables and directly derivable performance criteria and empirically derivable performance criteria. In the following we focus on criteria which can be derived directly from the floor plan geometry of an apartment. The list of criteria includes a variety of aspects ranging from accessibility, visibility, daylight illumination to energy consumption and common quantities.

**Accessibility**

By accessibility we understand the relationship between spaces. One important measure is how central a space in a configuration is. Based on the convex map we calculate the total number of steps needed to get from one convex space to all the others (depth). To be able to compare floor plans with different numbers of convex spaces, the measure is normalized by dividing by the number of convex spaces (Mean Depth). Additionally we calculate the number of convex spaces that have to be passed to get to each convex space, starting from the entrance (Step Depth from entrance). Based on this value, we can easily identify how far each space is away from the entrance (which represents interface to the exterior of an apartment).

Further measures are derived from a circulation diagram. The diagram is created based on the sight lines between doors. It represents the minimum space necessary to access every room in an apartment. To store this information in the DB and make it accessible for spatial statistics, the lines of movement are pixelated and stored as attributes of the analysis points. Additionally measures such circulation area per convex space, physical room and the entire flat are derived.

**Visibility**

By visibility we mean criteria that characterize the visual appearance of an apartment. For this we use the concept of isovists (Benedikt, 1979) to determine what can be seen from certain points (such as the entrance) and how the shape of the area that can be seen is structured (form factors of the isovist polygon such as area, compactness, occlusivity). The isovist is, of course, only an abstract concept which does not reflect the human field of vision, but enables us to make general statements when the direction of view is not known in advance. They have been widely used in psychological studies and have served as a useful predictor for the experiential qualities of space (Franz & Wiener, 2008; Stamps, 2009). For the analysis the floor plan is shown without doors. This is a common
convention in visibility analyses and helps to reveal the potential that the floor plan geometry offers.

Besides the calculation of the visual properties of single points in space, the relationships between these points are evaluated. Here we construct a visibility graph (Turner et al., 2001) and then derive the Visual Mean Depth from this graph, which indicates how easily a point in space can be seen from all the other points (the fewer changes in direction needed to see all the points, the more central a point is). Additionally we calculate the Visual Step Depth from entrance. This measure shows how often one has to move and change direction to see other points in the apartment.

To complement the visual properties related to the interior distribution of spatial elements, we developed a measure for evaluating views outside. To begin with a huge circle is created around the apartment and then the sections of the arc of this circle that are visible are calculated for each analysis point and totalled to ascertain the Exterior View Length. The measure is normalized from 0 to 1 (whereby 0 refers to no view at all and 1 refers to a 360° view without obstructions). The measures are calculated for two variants of window layouts, firstly the original arrangement of windows (OWL) and an idealized window layout whereby all exterior walls are assumed to be glazed (IWL). The latter makes it possible to assess the potential views that the arrangement of interior and exterior walls can offer.

**Daylight**
To describe the availability of daylight in an apartment we use the daylight factor (DF). The DF makes it possible to make general statements about the brightness of points in space without considering orientation. Since the availability of daylight is mainly influenced by the layout of windows, we examine the original window layout with the original ceiling height (OWL) as well as two idealized window layout variants (IWL). For the IWL we assume a fixed ceiling height of 2.7m and standardized window layouts (horizontal windows along the exterior walls) with a glazing ratio of 30% and 90% (IWL30, IWL90). The IWLs make it possible to minimize the influence of windows on daylight performance and to more adequately investigate the impact of the spatial arrangement of walls and doors. Furthermore we can use it to compare different floor plans. In addition to the standard values (average, minimum, maximum and standard deviation) for characterizing the larger spatial entities (convex spaces, rooms, apartment), we calculate the percentage of floor area with a DF below 2% as an indicator of the amount of well-illuminated floor area in a space.

**Energy Consumption**
To investigate the impact of floor plan geometry on energy consumption one needs to be aware that the main influencing factors are heat loss through exterior walls and solar gains through glazing. This is mainly influenced by the window layout and the orientation of the floor plan. To be able to compare different floor plans and investigate the impact of geometry on energy consumption, energy performance criteria (Heating/Cooling Demand) are calculated based on the same variants used for daylight analysis (OWL, IWL30, IWL90) with a standardized material and fixed location (Erfurt, Germany). In addition to the original orientation these variants are rotated in four directions (north, east, south and west) which produces energy consumption values for 15 variants per floor plan.

**Common Quantities**
This category summarizes common measures, such as total area, effective area, number of rooms/convex spaces, area of the single rooms/convex spaces, number of individually usable rooms (rooms without a fixed function, i.e. not rooms such as bathrooms or kitchens); length of exterior, neighbouring and interior walls. Based on these measures, we can derive secondary values such as building costs.

This list of measures is by no means exhaustive, but already illustrates how many criteria are influenced by the shape of the floor plan geometry. Figure 2 gives some examples. This shows how the de-
Figure 2
Different analysis diagrams for one apartment floor plan

- Step Depth from Entrance (Convex Map)
- Mean Depth (Convex Map)
- Circulation Diagram
- Visual Step Depth from Entrance (Visibility Graph)
- Visual Mean Depth (Visibility Graph)
- Isovist Area
- Isovist Compactness
- Isovist Occlusivity
- Exterior View Length (OWL)
- Exterior View Length (WLL)
- Daylight Factor (OWL)
- Daylight Factor (WLL30)
- Daylight Factor (WLL90)
- Heating Energy Demand (OWL, original orientation)
- Cooling Energy Demand (OWL, original orientation)
sign of a floor plan requires special care, along with a good understanding of how these parameters influence each other, and how they influence human behaviour. To ascertain meaningful correlations between the many different criteria, we need a large number of samples. The problem here is that manually evaluating the performance of a floor plan is very time-consuming. To assist in the process of data collection we propose a tool that aims to automate this process as far as possible. This tool is described in the next section.

AUTOMATED MODEL GENERATION AND PERFORMANCE EVALUATION

Different models are required to calculate the different performance criteria mentioned above. Accessibility analysis, for example, requires a convex map while visibility analysis needs a 2D plan with wall thickness. Daylight analysis requires a 3D model with thickened walls while Energy Analysis requires a simplified 3D model. Furthermore different variants of the model are required (OWL, IWL30, IWL90). To automate as much as possible of the process of model generation and performance evaluation we developed a tool called the "Floor Plan Analysis Toolbox" (FPAT). The FPAT was implemented in Grasshopper for Rhino (figure 3), because it provides a large number of geometric modelling tools and an ever-growing number of powerful analysis tools, making it easy to extend the FPAT in future.

Because original floor plan data is often stored using different media, levels of abstraction and symbolic languages, it is very hard to process them fully automatically without any human interaction. Instead, we rely on human input to interpret an ambiguous source of data but try to minimize the effort necessary to describe the apartment morphology. To achieve this as effectively as possible, the user only needs to create an abstract representation of a floor plan which is then used to generate the different models necessary for each analysis.

The level of abstraction is in our case a trade-off between accuracy and efficiency, which means that the chosen method is suitable for the majority of floor plans, with minimal loss of information and minimal time required for input. All the elements - exterior walls, neighboring walls, interior walls, convex space borders and doors - are represented by lines separated in different layers. Windows are represented by boxes that intersect the exterior walls. These elements form the basis for the generative process that creates all the necessary models (see figure 4).

In a first step, all walls are represented as a graph which is used to extract polygons that describe the physically closed rooms and convex spaces. Afterwards interconnections between convex spaces and physical rooms are calculated in order to create a topological graph and calculate its depth. For this we use a toolset called Decoding Spaces-Components. The space that is connected to the entrance (for step depth analysis) is recognized by door-lines that intersect with the neighboring or exterior walls.

To generate the 2D plan, the boundaries of the closed spaces are offset according to predefined thicknesses for each wall type (exterior, neighboring and interior walls). This model is used to calculate the usable area for each room/convex space. Using a flood-fill algorithm a grid of analysis points is generated by filling the 2D plan, starting from the centre of a convex space. For each of these points isovists, graph measures, external view length and the daylight factor is calculated. Isovist Properties, Visual Mean Depth and Visual Step Depth are calculated by using the Decoding Spaces Components.
For the daylight analysis, a 3D model is generated based on the thickened walls. Windows and doors are subtracted from the wall model using Boolean difference functions. The idealized window layouts (IWL30 and IWL90) are created by subtracting boxes, which are generated on basis of the exterior wall lines. Based on these 3D models, the daylight analysis is undertaken using the radiance-based simulation plugin DIVA.

For the energy analysis we generate a 3D zone model by extruding the room polygons and adding the floor and ceiling surface to obtain a closed Brep for every room. For the analysis we use the Energy-Plus-based energy calculation plugin ArchSim.

Finally we implemented a means of data export from Grasshopper to the database. The input geometry and the analysis results on the scale of the analysis grid, convex space and apartment are written into a CSV file. This is afterwards imported to the database. Nearly all the values and measures are generated automatically. Only the circulation diagram currently needs additional manual input. Depending on the size and complexity of a floor plan the whole process takes between 5 and 15 minutes. Afterwards some additional metadata can be added manually (such as year of construction, architect, geo-position, functions of rooms).

The major limitation of our current model is that it currently only facilitates single-storey apartments. On the one hand, this restrict the sample size of floor plans that can be examined but on the other allows us to use a simple and efficient user interface that does not require any complex modelling skills on the part of the user. Nevertheless the underlying generative and analytical model can be adapted in such a way that it could be transformed to analyse multi-storey apartments.

**PRELIMINARY TESTPHASE**

The FPAT has currently been tested in a seminar with 39 architecture students. In this preliminary test phase we collected 78 floor plans. Each student had to analyze two floor plans: one was the apartment he or she lived in and the other an apartment taken from a well-known floor plan manual (Schneider &
Heckmann 2011). Figure 5 shows the analysis diagrams for 8 different floor plans. The data was stored in the database via a web interface, making it possible to read the CSV-file from the FPAT. Various items of metadata were added manually via a web form. In the context of the seminar the data was used to explore the basic principles of knowledge discovery using databases. The collected data was accessed via MySQL queries and analyzed using simple statistical analysis (regression analysis). Students were able to investigate the influence of glazing area on daylight performance, the relationships between isovist properties and apartment size and between accessibility and visibility. The results of this seminar will be part of a further publication.

CONCLUSION & OUTLOOK
In this paper we presented a tool (FPAT) for enhancing the process of data collection for morphological studies on apartment floor plans. The FPAT makes it possible to calculate a multitude of directly calculable performance criteria from a simple floor plan sketch. In future we intend to extend the FPAT to easily input empirical data for space use (such as location of different items of furniture, most liked places, time spent in certain locations, movement patterns). This data is crucial for Environment-Behavior-Studies aimed at discovering relationships between directly calculable performance criteria and human behavior. To obtain a large amount of data on user experiences, a crowdsourcing approach could be promising. In terms of its technical implementation, a free-hand sketching approach combined with standardized building information models offers huge poten-
Besides benefitting academic research, the database can also be used by architects to search more flexibly for floor plans. These can be searched not just based on common quantities (such as number of rooms, size) but based on performance criteria. This opens up a completely new field of queries focusing on the qualities that a floor plan offers. It becomes possible to search for "floor plans with a visually central and naturally lit kitchen, where two rooms in the apartment offer a high degree of privacy" or "find apartments with three rooms, where two of them cannot be seen from the entrance".

Currently we are collaborating with the Chair of Usability from the Media Faculty at the Bauhaus-University Weimar to design an interface for interacting easily with the database. One of the central questions is how to deal with the vast amount of criteria in the search process.

ACKNOWLEDGEMENTS
This study was carried out as part of the FOGEB (Green Efficient Buildings) research project funded by the Thuringian Ministry for Economics, Labour and Technology and the European Social Funds (ESF).

REFERENCES
Hanson, J 1998, Decoding Homes and Houses, Cambridge University Press, Cambridge
Heckmann, O and Schneider, F (eds) 2011, Floor Plan Manual: Housing, Birkhäuser, Basel
Hillier, B, Hanson, J and Graham, H 1987, 'Ideas are in things: an application of the space syntax method to discovering house genotypes', Environment and Planning B: Planning and Design, 14, p. 363–385
Peponis, J 2012, 'Building Layouts As Cognitive Data: Purview And Purview Interface', Cognitive Critique, 6, pp. 11–51
Sevtsuk, A 2010, Path and place: a study of urban geometry and retail activity, Ph.D. Thesis, Massachusetts Institute of Technology
Steadman, P 1983, Architectural Morphology: An Introduction to the Geometry of Building Plans, Pion Ltd
Tanaka, JS 1987, "How big is big enough?": Sample size and goodness of fit in structural equation models with latent variables', Child Development, 58, pp. 134–146